REPORT DOCUMENTATION PAGE

Form Approved

OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
26-06-2014	Final	26 Mar 12 – 25 Mar 14

4. TITLE AND SUBTITLE

(124049) Robust Multimodal Cognitive Load Measurement

5a. CONTRACT NUMBER
FA23861214049

5b. GRANT NUMBER
Grant AOARD-124049

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)

Dr. Fang Chen

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
National ICT Australia (NICTA)
Level 5, 13 Garden Street
Eveleigh, NSW Sydney 2015
Australia

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD
UNIT 45002
APO AP 96338-5002

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/AFOSR/IOA(AOARD)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AOARD-124049

12. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution A: Approved for public release. Distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The objective of this project includes research of the fundamental issues related to the use of multiple input modalities and their fusion to enable robust and automatic cognitive load measurement (CLM) in the real world. Firstly, we carried out a further literature review on physiological measures of cognitive workload to include the recent advances of physiological measures of cognitive workload. In the meantime, we examined the use of various features (e.g. spectral and approximate entropies, wavelet-based complexity measures, correlation dimension, Hurst exponent) of electroencephalogram (EEG) signals to evaluate changes in working memory load during the performance of a cognitive task with varying difficulty/load levels. Eye based CLM was also studied. Eye activities such as pupillary response, blink, and eye movement (fixation and saccade) were investigated for CLM. We further investigated the linguistic and grammatical feature based CLM in this study and analyzed novel linguistic features as potential indices of cognitive load. Other modalities such as Galvanic skin response (GSR), face, and writing behavior were also extensively analyzed in indexing cognitive load levels. We also investigated the effect of stress on cognitive load. All together, we had carried out CLM study of multiple unobtrusive modalities, namely EEG, eye activity, linguistic and grammatical features, GSR, face, and writing behavior for CLM as well as emotion interference for CLM, in the past two-year period.

15. SUBJECT TERMS
Brain Science and Engineering, Computer and User Interface, Congnitive Modeling, Cognitive Modeling

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Brian Sells, Lt Col, USAF
U	U	U	SAR	13	**19b. TELEPHONE NUMBER** *(Include area code)* +81-3-5410-4409

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
16 JUL 2014	**Final**	**26-03-2012 to 25-03-2014**

4. TITLE AND SUBTITLE
(124049) Robust Multimodal Cognitive Load Measurement

5a. CONTRACT NUMBER
FA23861214049

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
Fang Chen

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
National ICT Australia (NICTA),Level 5, 13 Garden Street,Eveleigh, NSW Sydney 2015,Australia,NA,NA

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AOARD, UNIT 45002, APO, AP, 96338-5002

10. SPONSOR/MONITOR’S ACRONYM(S)
AOARD

11. SPONSOR/MONITOR’S REPORT NUMBER(S)
AOARD-124049

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The objective of this project includes research of the fundamental issues related to the use of multiple input modalities and their fusion to enable robust and automatic cognitive load measurement (CLM) in the real world. Firstly, we carried out a further literature review on physiological measures of cognitive workload to include the recent advances of physiological measures of cognitive workload. In the meantime, we examined the use of various features (e.g. spectral and approximate entropies, wavelet-based complexity measures, correlation dimension, Hurst exponent) of electroencephalogram (EEG) signals to evaluate changes in working memory load during the performance of a cognitive task with varying difficulty/load levels. Eye based CLM was also studied. Eye activities such as pupillary response, blink, and eye movement (fixation and saccade) were investigated for CLM. We further investigated the linguistic and grammatical feature based CLM in this study and analyzed novel linguistic features as potential indices of cognitive load. Other modalities such as Galvanic skin response (GSR), face, and writing behavior were also extensively analyzed in indexing cognitive load levels. We also investigated the effect of stress on cognitive load. All together, we had carried out CLM study of multiple unobtrusive modalities, namely EEG, eye activity, linguistic and grammatical features, GSR, face, and writing behavior for CLM as well as emotion interference for CLM, in the past two-year period.

15. SUBJECT TERMS
Brain Science and Engineering, Computer and User Interface, Congnitive Modeling, Cognitive Modeling

<table>
<tr><td colspan="3">16. SECURITY CLASSIFICATION OF:</td><td rowspan="2">17. LIMITATION OF ABSTRACT
Same as Report (SAR)</td><td rowspan="2">18. NUMBER OF PAGES
13</td><td rowspan="2">19a. NAME OF RESPONSIBLE PERSON</td></tr>
<tr><td>a. REPORT
unclassified</td><td>b. ABSTRACT
unclassified</td><td>c. THIS PAGE
unclassified</td></tr>
</table>

Final Report for AOARD Grant FA2386-12-1-4049

Robust Multimodal Cognitive Load Measurement (RMCLM)

March 26, 2014

Name of Principal Investigators (PI and Co-PIs): Fang Chen

- e-mail address : fang.chen@nicta.com.au
- Institution : National ICT Australia (NICTA)
- Mailing Address : Level 5, 13 Garden Street, Eveleigh NSW 2015, Australia
- Phone : +61 2 9376 2101
- Fax : +61 2 9376 2025

Period of Performance: March/26/2012 – March/25/2014

Abstract: This report summarizes the important research activities, study results and research accomplishments out of the RMCLM project in the past two-year period. The objective of this project includes research of the fundamental issues related to the use of multiple input modalities and their fusion to enable robust and automatic cognitive load measurement (CLM) in the real world. Firstly, we carried out a further literature review on physiological measures of cognitive workload to include the recent advances of physiological measures of cognitive workload. In the meantime, we examined the use of various features (e.g. spectral and approximate entropies, wavelet-based complexity measures, correlation dimension, Hurst exponent) of electroencephalogram (EEG) signals to evaluate changes in working memory load during the performance of a cognitive task with varying difficulty/load levels. Eye based CLM was also studied. Eye activities such as pupillary response, blink, and eye movement (fixation and saccade) were investigated for CLM. We further investigated the linguistic and grammatical feature based CLM in this study and analyzed novel linguistic features as potential indices of cognitive load. Other modalities such as Galvanic skin response (GSR), face, and writing behavior were also extensively analyzed in indexing cognitive load levels. We also investigated the effect of stress on cognitive load. All together, we had carried out CLM study of multiple unobtrusive modalities, namely EEG, eye activity, linguistic and grammatical features, GSR, face, and writing behavior for CLM as well as emotion interference for CLM, in the past two-year period.

List of Publications

a) Papers published in peer-reviewed journals

1. Khawaja, M. A., Chen, F., and Marcus, N., "**Measuring Cognitive Load using Linguistic Features - Implications for Usability Evaluation and Adaptive Interaction Design**", International Journal of Human-Computer Interaction, vol. 30, no. 5, pp. 343-368, 2014.
2. Hussain, M. S., Calvo, R. A., and Chen, F., "**Automatic cognitive load detecting from face, physiology, task performance and fusion during affective interference**", Interacting with Computers, vol. 25, no. 4, 2013. Elsevier.
3. Chen, S., and Epps, J., "**Blink Analysis for Cognitive Load Estimation: Towards Wearable Computing that Understands Your Current Task**", IEEE Pervasive Computing, vol.12, no.3, 56-65, 2013.
4. Zarjam, P., Epps, J., Chen, F., and Lovell, N. H., "**Estimating cognitive workload using wavelet entropy-based features during an arithmetic task**", Computers in Biology and Medicine, vol. 43, no.12, pp. 2186-2195, 2013. Elsevier.
5. Chen, F., Ruiz, N., Choi, E., Epps, J., Khawaja, A., Taib, R., Yin, B. and Wang, Y., "**Multimodal Behaviour and Interaction as Indicators of Cognitive Load**", ACM Transactions on Interactive Intelligent Systems, vol. 2, no. 4, article 22, December 2012.

6. Khawaja, M. A., Chen, F., Marcus, N., "**Analysis of Collaborative Communication for Linguistic Cues of Cognitive Load**", International Journal of Human Factors and Ergonomic Society, vol. 54, no 4. pp 518-529, August 2012.
7. Chen, S., Epps, J., "**Automatic Classification of Eye Activity for Cognitive Load Measurement with Emotion Interference**", Computer Methods and Programs in Biomedicine, 2012.

b) Papers published in peer-reviewed conference proceedings:

8. Chen, S., Epps, J., Chen, F., "**An Investigation of Pupil-based Cognitive Load Measurement with Low Cost Infrared Webcam under Light Reflex Interference**", Proc. EMBC 2013, Osaka, Japan, 2013.
9. Wang, W., Li, Z., Wang, Y., and Chen, F., "**Indexing cognitive workload based on pupillary response under luminance and emotional changes**", Proc. Int'l Conf. Intelligent User Interfaces (IUI), pp. 247-256, 2013.
10. Nourbakhsh, N., Wang, Y., and Chen, F., "**GSR and blink features for cognitive load classification**", Proc. IFIP Conference on Human-Computer Interaction (INTERACT), 2013.
11. Conway, D., Dick, I., Li, Z., Wang, Y., and Chen, F., "**The effect of stress on cognitive load measurement**", Proc. IFIP Conference on Human-Computer Interaction (INTERACT), 2013.
12. Yu, K., Epps, J., and Chen, F., "**Mental Workload Classification via Online Writing Features**", Proc. Int'l Conf. Document Analysis & Recognition (ICDAR) 2013.
13. Zarjam, P., Epps, J., Chen, F., and Lovell, N. H., "**Classification of Working Memory Load Using Wavelet Complexity Features of EEG Signals**", Lecture Notes in Computer Science, vol. 7664, pp. 692–699, Nov. 2012, Springer-Verlag Berlin.
14. Zarjam, P., Epps, J., Lovell, N. H., and Chen, F., "**Characterization of Memory Load in an Arithmetic Task using Non-Linear Analysis of EEG Signals**", Proc. of the 34th IEEE Engineering in Medicine and Biology Conference (EMBC'2012), pp. 3519-3522, California, USA, 2012.

Contents

1. Introduction

Historically, cognitive load has been measured using subjective self-rating scales (e.g. NASA TLX) and by performance scores, however these methods are post-hoc, are not feasible in all applications and are either subjective (self-rating) or not indicative of spare mental capacity (performance). There is a need for objective measures of cognitive load that are non-intrusive and objective, and have the potential to be determined in real time, i.e. measured continuously through the task.

This project has focused on multiple modalities, namely electroencephalogram (EEG), eye activity, GSR, face, linguistic and grammatical features, and writing behavior, for automatic cognitive load measurement.

2. Updated Literature Review

We carried out a further literature review on physiological measures of cognitive workload. The further investigation was focused on the recent advances of physiological measures of cognitive workload: eye movement, skin temperature, linguistic features, speech signals, EEG, Galvanic Skin Response (GSR), and pen input features.

3. EEG Based CLM

EEG is a noninvasive neuroimaging technique widely used for measuring cognitive workload, which offers high temporal resolution, ease of use, and a comparably low cost. We investigated different analysis method of electroencephalogram (EEG) signals to examine changes in working memory load during the performance of a cognitive task with varying difficulty levels.

3.1. Non-Linear Analysis of EEG Signals for CLM

Experiment: EEG signals were recorded during an arithmetic task while the induced load was varying in seven levels from very easy to extremely difficult. We studied six male participants, between the ages of 24-30 years. They were right-handed and had normal or corrected to normal eyesight and gave written informed consent, in accordance with human research ethics guidelines. We designed an addition task with seven levels of difficulty, starting from one digit addition (very low) to multi-digit addition (extremely difficult) as shown in Table I.

TABLE I.
TASK DIFFICULTY LEVEL DETAILS.

Task level	Number of digits	Example
Very low (L1)	1&2 digit numbers	45+2
Low (L2)	1&2 digit numbers with 1 carry	54+9
Medium (L3)	2 digit numbers with 1 carry	67+42
Medium-High (L4)	2 digit numbers with 2 carries	39+65
High (L5)	2&3 digit numbers with 1 carry	377+32
Very high (L6)	2&3 digit numbers with 2 carries	76+347
Extremely high (L7)	3 digit numbers with 3 carries	983+748

The EEG signals were recorded from 32 channels mounted in an elastic cap, according to the extended international 10-20 system using an Active Two acquisition system. The experiment was conducted under controlled conditions in an electrically isolated laboratory, with a

minimum distance of five meters from power sources to the experiment desk and under natural illumination.The EEG signals were analyzed using three different non-linear/dynamic measures. They were correlation dimension (CD), Hurst exponent (HE) and approximate entropy (ApEn).

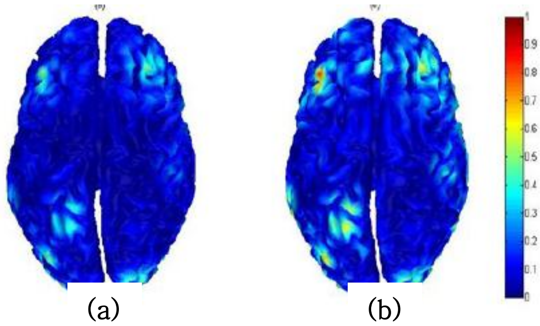

Figure 1. The source maps of two load levels for subject 1; (a) the lowest load (L1), and (b) the most difficult load (L7). Both load levels influence the similar regions more or less but the degree of activation increased as the load level increased.

Results and Discussion: Experimental results show that the values of the measures extracted from the delta frequency band of signals acquired from the frontal and occipital lobes of the brain vary in accordance with the task difficulty level induced (see Figure 1). The values of the correlation dimension increased as the task difficulty increased, showing a rise in complexity of the EEG signals, while the values of the Hurst exponent and approximate entropy decreased as task difficulty increased, indicating more regularity and predictability in the signals.

3.2. Wavelet Complexity Features of EEG Signals for CLM

Experiment: In this study, the use of wavelet-based complexity measures of EEG signals were investigated to evaluate changes in working memory load during the performance of a cognitive task with varying difficulty/load levels. EEG signals were acquired from twelve healthy male subjects; postgraduate students aged between 24-30 years. In the experiment, the participants were asked to do an arithmetic task (an addition task with varying difficulty level, see Table I).

The subjects' EEG signals were recorded using an Active Two system. Each recording contained 32 EEG channels mounted in an elastic cap, according to the extended international 10 - 20 system. A linked earlobe reference was used and impedance was kept under 5 kΩ. The EEG signals were passed through a band-pass filter with cut-off frequencies of 0.1 - 100 Hz and were recorded at an fs =256 Hz sampling rate. To select the epochs which contained minimal EMG artifact, each recording was judged by visual inspection. As a result, 70 seconds (out of 90 seconds of each task level recording) for each subject was considered. This portion of the recordings included EOG and ECG artifacts, which were not removed.

Extracted signals were analyzed using wavelet based complexity measures. The wavelet complexity measures associate with four entropic measures: that is Shannon, Tsallis, Escort-Tsallis and Renyi entropies.

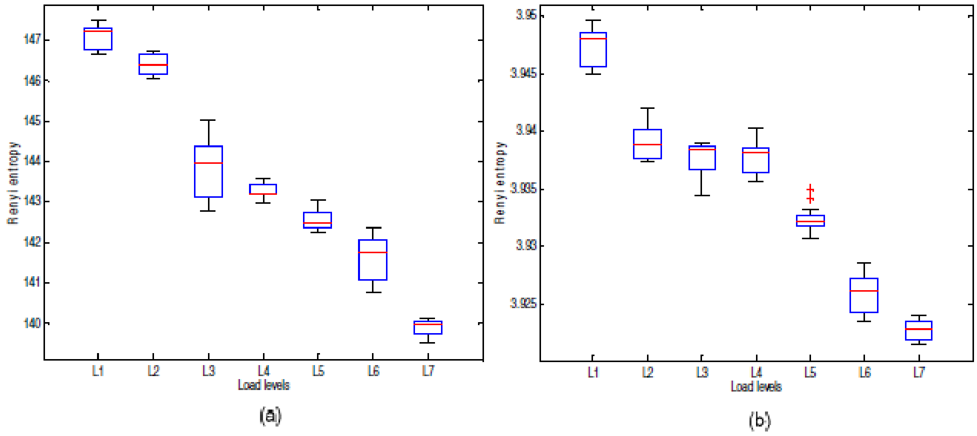

Figure 2. The Renyi entropy variations for (a) *q*=0.9, (b) *q*=0.1 with the load levels, for channel F7 of subject 1. On each box, the red mark is the median; the edges of the box are the 25th and the 75th percentiles.

Results and Discussion: As an example, Figure 2 shows the median of the extracted H_{RE} from the frontal channels in scale 5, for channel F7 of subject 1, for two extreme values of *q* (entropic index); (a) *q*= 0.9, (b) *q*= 0.1, in the delta frequency band. As shown, the median of the extracted H_{RE} are able to distinguish the seven task loads better with *q* closer to 1, as it consistently reveals a decreasing median with increasing task load.

The experimental results demonstrated good discrimination among seven load levels imposed on the working memory with a classification rate of up to 96% using signals recorded from the frontal lobe of the brain. The extracted measures' values show a consistent decrease in the selected channels in two frontal and occipital lobes, as the memory load increases, indicating the EEGs disorder declines while the complexity grows. This illustrates that the brain behaves in a more organized manner characterized by more order and maximal complexity when dealing with higher load levels. The growing complexity can also reflect the higher activation of neural networks involved, as the task load increases.

3.3. Entropy Based Features of EEG Signals for CLM

Experiment: In this study, we investigated the use of entropy-based features (spectral and approximate entropies) of recorded EEG signals to characterize mental load when performing a cognitive task. The participants' EEG signals were recorded using the same method as in the study of Non-Linear Analysis of EEG Signals for CLM and Wavelet Complexity Features of EEG Signals for CLM (six participants were involved in the experiment).

The recorded EEG signals were analysed using following methods: 1) EEG signal source localization using the minimum norm estimate algorithm, 2) sub-band filtering by Discrete Wavelet Transform (DWT), 3) entropy-based feature extraction from the EEG signals.

Results and Discussion: The experimental results demonstrated that the spectral entropy is a good discriminator of mental load level and decreases consistently in accordance with the increased load. The extracted approximate entropy quantifies the irregularity of the EEGs, indicating a decrease in irregularity as the load increases. We also perform EEG source estimation to choose a smaller subset of EEG channels which make the most contribution in the load level discrimination. We conclude that the entropy-based features are capable of measuring the imposed mental load from the selected channels in two brain regions. This may demonstrate that the brain behaves in a more regular or focused manner when dealing with higher task loads. The efficacy of entropy-based features across frequency subbands was also analyzed in this study.

4. Eye-Based CLM

Eye activity has advantages in CLM. For example, eye activity is more ubiquitous than other modalities; pupillary response and eye blink have been shown to correlate with both visual and aural cognitive tasks; eye activity data collection is less intrusive than other physiological signal data collection. Eye-based CLM is a popular physiological index of cognitive workload that can be used for design and evaluation of adaptive interface in various areas of human-computer interaction (HCI) research.

Eye-based automatic CLM was studied in our research. Three types of eye activity were investigated: pupillary response, blink, and eye movement (fixation and saccade). Eye activity features were investigated in the presence of emotion interference, which is a source of undesirable variability, to determine the susceptibility of CLM systems to other factors.

4.1. Eye Activity for CLM with Emotion Interference

Experiment: In this study, cognitive load was induced using arithmetic tasks, and the difficulty level was controlled by the number of carries and digits. Emotional interference corresponding to different arousal and valence levels was induced by showing International Affective Picture System (IAPS) images in the task background. The experiment was adapted from those using pupillary response for measuring cognitive load with arithmetic tasks and for measuring arousal with IAPS images.

The participants comprised seven females and eight males, aged 20–48. A total of 82 recordings were obtained from each participant, including 60 samples with both cognitive load and emotion factors, 10 samples with only the cognitive load factor and 12 samples with only the emotion factor. The signal length of each sample was 14 s, during which four task stimuli were systematically presented and time stamped. Figure 3 shows the time line for each task.

Figure 3. Time line for each task. Each task comprises focusing, image viewing, reading and calculating four addends sequentially, selecting an answer and subjective rating of both task difficulty and emotion.

Results and Discussion: ANOVA test results and multiple regression results revealed important implications for using eye activity for CLM. Pupil size and blink number increased with more difficult tasks, which perfectly matches the literature. Pupil size also increased with higher arousal images regardless of valence, which is also consistent with studies of pupil dilation using visual and auditory stimuli. However, pupil size increased with images of positive valence when a task goal was presented in this study, as the p value was close to 0.05.

The new finding here was that some eye activity feature patterns (notably pupil dilation and blink) for the cognitive load levels were not significantly altered with or without arousal factor in the task-goal driven situation. In contrast, the patterns of features for the arousal level seemed weakened in the pupillary response when cognitive load was induced and there was

no arousal effect on the features of blink, fixation and saccade. This result suggests the dominance of cognitive load over emotion in eye features during task performance.

4.2. Pupil-Based CLM with Low Cost Infrared Webcam

Experiment: In this study, we investigate the validity of CLM with (i) pupil light reflex in a less controlled luminance background; (ii) a low-cost infrared (IR) webcam for the pupillary response in a controlled luminance background. We employed a mental arithmetic task but in three different conditions, (i) constant gray background with FaceLab4 - a remote eye tracker ('gray_eye_tracker', 15 participants), where there is little PLR due to global luminance and minimum pupil light reflex (PLR) due to gaze shift; (ii) image background with the remote eye tracker ('image_eye_tracker', 15 participants), where both the PLR due to different global luminance and the PLR due to gaze shift occur; and (iii) constant gray background with IR webcams ('gray_webcam', 22 participants) for pupillary response recording.

Results and Discussion: ANOVA results showed that with an appropriate baseline selection and subtraction, the light reflex was significantly reduced, suggesting the possibility of less constrained practical applications of pupil-based CLM. Compared with the pupillary response from a commercial remote eye tracker, a low-cost IR webcam achieved a similar pupillary response pattern and no significant difference was found between the two devices in terms of cognitive load measurement across five induced load levels.

In another similar experiment, it was found that blink features were more sensitive to perceptual load than cognitive load and task transition.

4.3. Pupillary Response for CLM under Luminance and Emotional Changes

Experiment: This research investigated pupillary response as a cognitive workload measure under the influence of luminance and emotional changes. Two experiments were conducted: 1) Thirteen 24-to-46-year-old male subjects were invited to perform arithmetic tasks under different luminance conditions on the monitor screen. The arithmetic tasks were designed to have four difficulty levels, and there were four levels of background brightness, resulting in 16 different task types in total. 2) Twelve 24-to-35-year-old male participants perform arithmetic tasks under the changes of luminance condition and emotional arousal.

Results and Discussion: The proposed mean-difference feature and its extension (Haar-like features) have been demonstrated to effectively characterize physiological responses of cognitive workload under luminance and emotional changes. Boosting based feature selection and classification has been employed to robustly classify workload even under the influence of those noisy factors. The proposed technique could be applied to various applications involving cognitive workload evaluation under complex environments.

4.4. GSR and Blink Features for CL Classification

Experiment: This study aimed to analyze GSR and blink features for the classification of CL. The data was collected from thirteen healthy 24 to 35-year-old volunteers. The experiment included 8 arithmetic tasks with 4 difficulty levels. Each subject performed two trials of each task level and the whole eight trials were performed in a randomized order. In each task four numbers were shown one by one, each for three seconds. Subjects were supposed to add-up these four numbers and select (by clicking the mouse using their right hand) the correct answer from three numbers which were next presented on the screen.

Results and Discussion: Two GSR and two blink features were calculated for each task: 1) accumulative GSR (summation of GSR values over task time), 2) GSR power spectrum (frequency power), 3) blink number (number of blinks in the task), 4) blink rate (number of blinks in the task divided by task time). Support vector machines (SVM) and Naïve Bayes classifiers were applied for cognitive load classification. Obtained results show that the studied features of blink and GSR can reasonably discriminate workload levels and combining features

of the two modalities improves the accuracy of cognitive load classification. We have achieved around 75% for binary classification and more than 50% for four-class classification.

5. Face Modality and Fusion for CLM

This research investigated automatic CLM from facial features, physiology and task performance under affective interference.

Experiment: 20 participants (11 males and 9 females, whose age ranged from 22 to 48) were recruited to solve mental arithmetic tasks with emotional stimuli in the background as shown in Figure 4. The number of digits for addition and the number of carries produced by addition determined the level of task difficulty. The images were selected from IAPS according to the normative ratings of valence and arousal to induce emotions, and they were presented based on the affective circumplex model.

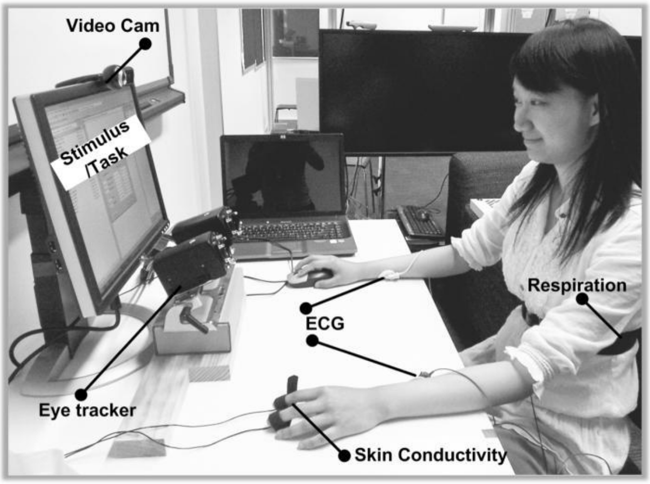

Figure 4. Sample experimental setup and sensors.

Results and Discussion: Two types of image-based features (geometric and chromatic) were extracted from face video. Statistical features such as mean, median, standard deviation, minima and maxima were computed. Different actions of the user’s head, such as tilting, nodding and leaning forward or backward, were calculated and detected by subtracting the values of the last frame of the time window from the first values. A total of 115 unique features were extracted from the video (59 from geometric and 56 from chromatic). Pupil and eye movement features were also extracted from the data collected with the eye tracker and webcam. Performance features were extracted from experiment logs. Results indicated that the face modality for cognitive load detection was more accurate under affective interference, whereas physiology and task performance were more accurate without the affective interference. Multimodal fusion improved detection accuracies, but it was less accurate under affective interferences. More specifically, the accuracy decreased with increasing intensity of emotional arousal.

6. Linguistic and Grammatical Features for CLM

Linguistic and grammatical features may be extracted from users’ spoken language and analysed for patterns indicating high cognitive load. These features may include speech pauses, self-corrections, repetitions, response latency, and language usage, for example, use of different word categories and parts of speech, such as nouns and pronouns, and grammatical structures. Such features may be collected from users’ spoken or written

language and are highly unobtrusive. Linguistic features have been regarded as indices of high cognitive load.

Experiment: This research studied 33 members of bushfire management teams working collaboratively in computerized incident control rooms and involved in complex bushfire management tasks. The team members carried out 10 tasks, each about 5 hr in duration, in four states of Australia, including New South Wales, Victoria, Tasmania, and Queensland.

The participants' communication was analyzed for some novel linguistic features as potential indices of cognitive load, which included sentence length, use of agreement and disagreement phrases, and use of personal pronouns, including both singular and plural pronoun types.

Results and Discussion: The experimental results confirmed that while working collaboratively and performing high-cognitive load tasks, people speak more with other team members to manage and share the high task complexity. The results showed that participants, especially those working in a collaborative team environment, consistently use singular pronouns and plural pronouns differently in different task load situations. Specifically, they used significantly more singular pronouns for low-load tasks than for high-load tasks; that is, the lower the cognitive demand, the greater use of singular pronouns. In contrast, they used significantly more plural pronouns for highload tasks than for low-load tasks; that is, the higher the cognitive load, the greater use of plural pronouns. These results support the notion that people actually collaborate and coordinate tasks more with each other during highly complex real-world tasks.

In another similar experiment, it was found that linguistic category features (e.g. negative emotions, positive emotions, agreement words), and language complexity measures (e.g. lexical density, complex word ratio, Gunning Fog Index) were significantly different under different CL levels.

7. Writing Features for CLM

Mental workload is an important factor during writing, which may affect the writing efficiency and user experience. This research aimed at a method to classify the mental workload levels during writing process, via examination of online writing features.

Experiment: In this study, we adapted experiments from Ransdell and Levy's experimental design, which is easy to deploy and convenient to induce different mental load levels. Specifically, one or more words were displayed on the task interface (Figure 5) and then disappeared after specified time. The time to display the words varied from one second to 2.5 seconds depending on the number of given words. The participants were required to remember the given words after they disappeared, compose a sentence with the given words and write the sentence down on a writing space of 20cm×10cm on the WACOM DTZ-1200W tablet with a digital pen. 20 university students were recruited.

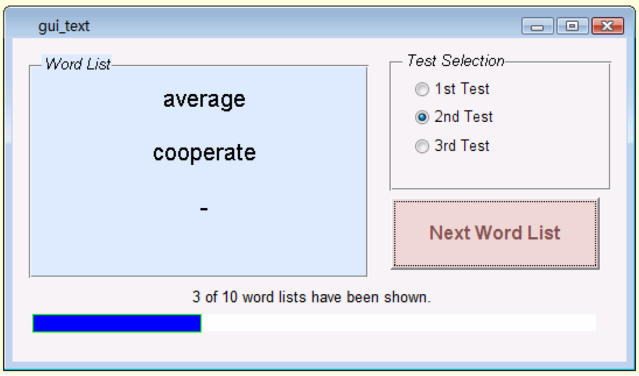

Figure 5. The task interface for the given words.

Results and Discussion: At the first stage, a curvature tracking method was applied to the handwriting script, to examine the curvature for individual writing points. Then a selection process allocated writing points into subsets, each corresponding to one curvature span. The second stage extracted velocity features, used to characterize mental workload, from points in each curvature span. A Parzen-window classifier was applied on velocity features from each curvature span. The classification decisions from individual classifiers were fused with a selective voting scheme for the overall mental workload classification decision. The fused curvature-dependent mental load classifier has been demonstrated to be feasible for mental workload classification.

8. The Effect of Stress on CLM

Stress has been shown to effect both the sympathetic and parasympathetic nervous systems and, in its more extreme states, results in large changes to physiological function that may well obscure the relationship between a physiological indicator and CL. Furthermore, stress may, in some circumstances, be a confounding factor for CL in that changes in CL may correlate with changes in stress levels. This research studied the effect of stress on CL measurement using GSR.

Experiment: 11 male students and employees (24-49 years' old, ten right handed and one left-handed) took part in the experiment. The experiment consisted of a within- subjects, six-way factorial design. There were math questions of three difficulty levels (low, medium and high) administered under two different stress conditions: 'no-stress' and 'stress'. The experiment utilized feelings of lack of control, task failure and social- evaluation to induce stress.

Results and Discussion: Without the impact of stress, it appeared that an increase in CL (induced by increasing the difficulty of tasks given to test subjects) resulted in an increase in mean GSR value. This relationship was, however, obfuscated when test subjects experienced fluctuating levels of stress. GSR may still be useful as an index for CL even when stress was a confounding factor, if we considered peak based features extracted from the GSR signal other than the mean value. Both peak frequency in the signal and peak durations were negatively correlated to task difficulty and hence CL. These features could possibly be used to dissociate CL from stress and develop a stress-agnostic method of CL classification.

9. Conclusions and Future Work

This research carried out CLM study of various unobtrusive modalities: EEG, eye activity, face, linguistic features, and writing behaviors for CLM.

In the EEG based CLM, we examined the use of various features (e.g. spectral and approximate entropies, wavelet-based complexity measures, correlation dimension, Hurst exponent) of EEG signals to evaluate changes in working memory load during the performance of a cognitive task with varying difficulty/load levels. Experimental results showed that EEG may be more reliable than self-rating, and capable of distinguishing seven load levels induced under controlled conditions with accuracies exceeding 94%.

In the eye based CLM, three types of eye activity were investigated: pupillary response, blink, and eye movement (fixation and saccade). Results from experiments combining arithmetic-based tasks and affective image stimuli demonstrated that arousal effects were dominated by cognitive load during task execution.

In the facial modality based CLM, two types of image-based features (geometric and chromatic) were extracted from face video. Other features such as statistical features and actions of the user’s head were extracted. Results indicated that the face modality for CLM was more accurate under affective interference. Multimodal fusion improved detection accuracies, but it was less accurate under affective interferences.

The linguistic feature based CLM was also investigated in this study. Some novel linguistic features were analyzed as potential indices of cognitive load. Results showed that with high load, people spoke more and used longer sentences, used more words that indicated disagreement with other team members, and exhibited increased use of plural personal pronouns and decreased use of singular pronouns. It was also found that linguistic category features (e.g. negative emotions, positive emotions, agreement words), and language complexity measures (e.g. lexical density, complex word ratio, Gunning Fog Index) were significantly different under different CL levels.

In the writing feature analysis for CLM, the fused curvature-dependent mental load classifier has been demonstrated to be feasible for mental workload classification.

We also investigated the effect of stress on CL.

Future work will include analyzing the cognitive workload based on understanding the contextual task characteristics and user behavior in interaction which can benefit the measurement of cognitive load and development of intelligent systems to aid user task management. The direct and continuous observations of individual tasks via eye activity as well as other physiological measurements will be investigated in the future work.